



SEQUENCE LISTING

(1) GENERAL INFORMATION:

- 5 (i) APPLICANT: Skatrud, Paul L.
Peery, Robert B.
de Waard, Maarten
- 10 (ii) TITLE OF INVENTION: Multiple Drug Resistance Gene atrD of
Aspergillus Nidulans
- 15 (iii) NUMBER OF SEQUENCES: 3
- (iv) CORRESPONDENCE ADDRESS:
20 (A) ADDRESSEE: Eli Lilly and Company
(B) STREET: Lilly Corporate Center
(C) CITY: Indianapolis
(D) STATE: Indiana
(E) COUNTRY: U.S.
(F) ZIP: 46285
- 25 (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- 30 (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:
- 35 (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Webster, Thomas D.
(B) REGISTRATION NUMBER: 39,872
(C) REFERENCE/DOCKET NUMBER: X-11766
- 40 (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 317-276-3334
(B) TELEFAX: 317-276-2763
- 45 (2) INFORMATION FOR SEQ ID NO:1:
- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4002 base pairs
(B) TYPE: nucleic acid
50 (C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- 55 (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- 60 (ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 1..4002
- 65 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

X-11766

17

	ATG TCC CCG CTA GAG ACA AAT CCC CTT TCG CCA GAG ACT GCT ATG CGC	48
	Met Ser Pro Leu Glu Thr Asn Pro Leu Ser Pro Glu Thr Ala Met Arg	
	1 5 10 15	
5	GAA CCT GCT GAG ACT TCA ACG ACG GAG GAG CAA GCT TCT ACA CCA CAC	96
	Glu Pro Ala Glu Thr Ser Thr Thr Glu Glu Gln Ala Ser Thr Pro His	
	20 25 30	
10	GCT GCG GAC GAG AAG AAA ATC CTC AGC GAC CTC TCG GCT CCA TCT AGT	144
	Ala Ala Asp Glu Lys Lys Ile Leu Ser Asp Leu Ser Ala Pro Ser Ser	
	35 40 45	
15	ACT ACA GCA ACC CCC GCA GAC AAG GAG CAC CGT CCT AAA TCG TCG TCC	192
	Thr Thr Ala Thr Pro Ala Asp Lys Glu His Arg Pro Lys Ser Ser Ser	
	50 55 60	
20	AGC AAT AAT GCG GTC TCG GTC AAC GAA GTC GAT GCG CTT ATT GCG CAC	240
	Ser Asn Asn Ala Val Ser Val Asn Glu Val Asp Ala Leu Ile Ala His	
	65 70 75 80	
25	CTG CCA GAA GAC GAG AGG CAG GTC TTG AAG ACG CAG CTG GAG GAG ATC	288
	Leu Pro Glu Asp Glu Arg Gln Val Leu Lys Thr Gln Leu Glu Ile	
	85 90 95	
30	AAA GTA AAC ATC TCC TTC GGT CTC TGG CGG TAT GCA ACA AAG ATG	336
	Lys Val Asn Ile Ser Phe Phe Gly Leu Trp Arg Tyr Ala Thr Lys Met	
	100 105 110	
35	GAT ATA CTT ATC ATG GTA ATC AGT ACA ATC TGT GCC ATT GCT GCC GCG	384
	Asp Ile Leu Ile Met Val Ile Ser Thr Thr Ile Cys Ala Ile Ala Ala	
	115 120 125	
40	TCG ACT TTC CAG AGG ATA ATG TTA TAT CAA ATC TCG TAC GAC GAG TTC	432
	Ser Thr Phe Gln Arg Ile Met Leu Tyr Gln Ile Ser Tyr Asp Glu Phe	
	130 135 140	
45	TAT GAT GAA TTG ACC AAG AAC GTA CTG TAC TTC GTA TAC CTC GGT ATC	480
	Tyr Asp Glu Leu Thr Lys Asn Val Leu Tyr Phe Val Tyr Leu Gly Ile	
	145 150 155 160	
50	GCG GAG TTT GTC ACT GTC TAT GTT AGT ACT GTT GGC TTC ATC TAT ACC	528
	Gly Glu Phe Val Thr Val Tyr Val Ser Thr Val Val Gly Phe Ile Tyr Thr	
	165 170 175	
55	GGA GAA CAC GCC ACG CAG AAG ATC CGC GAG TAT TAC CTT GAG TCT ATC	576
	Gly Glu His Ala Thr Gln Lys Ile Arg Glu Tyr Tyr Leu Glu Ser Ile	
	180 185 190	
60	CTG CGC CAG AAC ATT GGC TAT TTT GAT AAA CTC GGT GCC GGG GAA GTG	624
	Leu Arg Gln Asn Ile Gly Tyr Phe Asp Lys Leu Gly Ala Gly Glu Val	
	195 200 205	
65	ACC ACC CGT ATA ACA GCC GAT ACA AAC CTT ATC CAG GAT GGC ATT TCG	672
	Thr Thr Arg Ile Thr Ala Asp Thr Asn Leu Ile Gln Asp Gly Ile Ser	
	210 215 220	
70	GAG AAG GTC GGT CTC ACT TTG ACT GCC CTG GCG ACA TTC GTG ACA GCA	720
	Glu Lys Val Gly Leu Thr Leu Thr Ala Leu Ala Thr Phe Val Thr Ala	
	225 230 235 240	
75	TTC ATT ATC GCC TAC GTC AAA TAC TGS AAG TTG GCT CTA ATT TGC AGC	768
	Phe Ile Ile Ala Tyr Val Lys Tyr Trp Lys Leu Ala Leu Ile Cys Ser	
	245 250 255	
80	TCA ACA ATT GTG GCC CTC GTT CTC ACC ATG GGC GGT GGT TCT CAG TTT	816

	Ser Thr Ile Val Ala Leu Val Leu Thr Met Gly Gly Gly Ser Gln Phe	
	260 265	
5	ATC ATC AAG TAC AGC AAA AAG TCG CTT GAC AGC TAC GGT GCA GGC GGC Ile Ile Lys Tyr Ser Lys Lys Ser Leu Asp Ser Tyr Gly Ala Gly Gly	864
	275 280 285	
10	ACT GTT GCG GAA GAG GTC ATC AGC TCC ATC AGA AAT GCC ACA GCG TTT Thr Val Ala Glu Glu Val Ile Ser Ser Ile Arg Asn Ala Thr Ala Phe	912
	290 295 300	
	GGC ACC CAA GAC AAG CTT GCG AAG CAG TAT GAG GTC CAC TTA GAC GAA Gly Thr Lys Asp Lys Leu Ala Lys Gln Tyr Ile Val Met Gly Phe Met Ile	960
15	305 310 315 320 325	
	GCT GAG AAA TGG GGA ACA AAG AAC CAG ATT GTC ATG GGT TTC ATG ATT Ala Glu Lys Trp Gly Thr Lys Asn Gln Ile Val Met Gly Phe Met Ile	1008
	330 335	
20	GGC GCC ATG TTT GGC CTT ATG TAC TCG AAC TAC GGT CTT GGC TTC TGG Gly Ala Met Phe Gly Leu Met Tyr Ser Asn Tyr Gly Leu Gly Phe Trp	1056
	340 345	
25	ATG GGT TCT CGT TTC CTG GTA GAT GGT GCA GTC GAT GTG GGT GAT ATT Met Gly Ser Arg Phe Leu Val Asp Gly Ala Val Asp Gly Asp Ile	1104
	355 360	
30	CTC ACA GTT CTC ATG GCC ATC TTG ATC GGA TCG TTC TCC TTG GGG AAC Leu Thr Val Leu Met Ala Ile Leu Ile Gly Ser Ser Leu Gly Asn	1152
	370 375 380	
	GTT AGT CCA AAT GCT CAA GCA TTT ACA AAC GCT GTG GCC GCG GCC GCA Val Ser Pro Asn Ala Gln Ala Phe Thr Asn Ala Val Ala Ala Ala	1200
35	385 390 395 400 405	
	AAG ATA TTT GGA ACG ATC GAT CGC CAG TCC CCA TTA GAT CCA TAT TCG Lys Ile Phe Gly Thr Ile Asp Arg Gln Ser Pro Leu Asp Pro Tyr Ser	1248
	410 415	
40	AAC GAA GGG AAG ACG CTC GAC CAT TTT GAG GGC CAC ATT GAG TTA CGC Asn Glu Gly Lys Thr Leu Asp His Phe Glu Gly His Ile Glu Leu Arg	1296
	420 425	
45	AAT GTC AAG CAT ATT TAC CCA TCT AGA CCC GAG GTC ACC GTC ATG GAG Asn Val Lys His Ile Tyr Pro Ser Arg Pro Glu Val Thr Val Met Glu	1344
	435 440 445	
	GAT GTT TCT CTG TCA ATG CCC GCT GGA AAA ACA ACC GCT TTA GTC GGC Asp Val Ser Leu Ser Met Pro Ala Gly Lys Thr Thr Ala Leu Val Gly	1392
50	450 455 460 465	
	CCC TCT GGC TCT GGA AAA AGT ACG GTG GTC GGC TTG GTT GAG CGA TTC Pro Ser Gly Ser Gly Lys Ser Thr Val Val Gly Leu Val Glu Arg Phe	1440
55	465 470 475 480 485	
	TAC ATG CCT GTT GCG GGT ACG GTT TTG CTG GAT GGC CAT GAC ATC AAG Tyr Met Pro Val Arg Gly Thr Val Leu Leu Asp Gly His Asp Ile Lys	1488
	485 490 495	
60	GAC CTC AAT CTC CGC TGG CTT CGC CAA CAG ATC TCT TTG GTT AGC CAG Asp Leu Asn Leu Arg Trp Leu Arg Gln Gln Ile Ser Leu Val Ser Gln	1536
	500 505 510	
65	GAG CCT GTT CTT TTT GGC ACG ACG ATT TAT AAG AAT ATT AGS CAC GGT Glu Pro Val Leu Phe Gly Thr Thr Ile Tyr Lys Asn Ile Arg His Gly	1584

	515	520	525	
5	CTC ATC GGC ACA AAG TAC GAG AAT GAA TCC GAG AAG GTC CGG GAA Leu Ile Glu Thr Lys Tyr Glu Asn Glu Ser Glu Asp Lys Val Arg Glu 530 535 540			1632
10	CTC ATC GAG AAC GCG GCA AAA ATG GCG AAT GCT CAT GAC TTT ATT ACT Leu Ile Glu Asn Ala Lys Met Ala Asn Ala His Asp Phe Ile Thr 545 550 555			1680
15	GCC TTG CCT GAA GGT TAT GAG ACC AAT GTT GGG CAG CGT GGC TTT CTC Ala Leu Pro Glu Gly Tyr Glu Thr Asn Val Gly Gln Arg Gly Phe Leu 565 570 575			1728
20	CTT TCA GGT GGC CAG AAA CAG CGC ATT GCA ATC GCC CGT GCC GTT GTT Leu Ser Gly Gly Gln Lys Gln Arg Ile Ala Ile Ala Arg Ala Val Val 580 585 590			1776
25	AGT GAC CCA AAA ATC CTG CTC CTG GAT GAA GCT ACT TCG GCC TTG GAC Ser Asp Pro Lys Ile Leu Leu Leu Asp Glu Glu Thr Ser Ala Leu Asp 595 600 605			1824
30	ACA AAA TCC GAA GGC GTG GTT CAA GCA GCT TTG GAG AGG GCA GCT GAA Thr Lys Ser Glu Gly Val Val Gln Ala Ala Leu Glu Ala Ala Glu 610 615 620			1872
35	GCG CGA ACT ACT ATT GTG ATC GCT CAT CGC CTT TCC ACG ATC AAA ACG Gly Arg Thr Thr Ile Val Ile Ala His Arg Leu Ser Thr Ile Lys Thr 625 630 635 640			1920
40	GCG CAC AAC ATT GTG GTT CTG GTC AAT GGC AAA ATT GCT GAA CAA GGA Ala His Asn Ile Val Val Leu Val Asn Gly Lys Ile Ala Glu Gln Gly 645 650 655			1968
45	ACT CAC GAT GAA TTG GTT GAC CGC GGA GGC GCT TAT CGC AAA CTT GTG Thr His Asp Glu Leu Val Asp Arg Gly Gly Ala Tyr Arg Lys Leu Val 660 665 670			2016
50	GAG GCT CAA CGT ATC AAT GAA CAG AAG GAA GCT GAC GCC TTG GAG GAC Glu Ala Gln Arg Ile Asn Glu Gln Lys Glu Ala Asp Ala Leu Glu Asp 675 680 685			2064
55	GCC GAC GCT GAG GAT CTC ACG AAT GCA GAT ATT GCC AAA ATC AAA ACT Ala Asp Ala Glu Asp Leu Thr Thr Asn Ala Asp Ile Ala Lys Ile Lys Thr 690 695 700			2112
60	GCG TCA AGC GCA TCA TCC GAT CTC GAC GGA AAA CCC ACA ACC ATT GAC Ala Ser Ser Ala Ser Ser Asp Leu Asp Gly Lys Pro Thr Thr Ile Asp 705 710 715 720			2160
65	GCG ACG GGC ACC CAC AAG TCT GTT TCC AGC GCG ATT CTT TCT AAA AGA Arg Thr Gly Thr His Lys Ser Val Ser Ser Ala Ile Leu Ser Lys Arg 725 730 735			2208
70	CCC CCC GAA ACA ACT CCG AAA TAC TCA TTA TGG ACG CTG CTC AAA TTT Pro Pro Glu Thr Thr Pro Lys Tyr Ser Leu Trp Thr Leu Leu Lys Phe 740 745 750			2256
75	GTT GCT TCC TTC AAC CGC CCT GAA ATC CCG TAC ATG CTC ATC GGT CTT Val Ala Ser Phe Asn Arg Pro Glu Ile Pro Tyr Met Leu Ser Gly Leu 755 760 765			2304
80	GTC TTC TCA GTG TTA GCT GGT GGT GGC CAA CCC ACG CAA GCA GTG CTA Val Phe Ser Val Leu Ala Gly Gly Gln Pro Thr Gln Ala Val Leu 770 775 780			2352

[illegible]

	TCC GCG GGC ACC GTC TTT TCC TTT GCA CCA GAC ATG GGC AAG GCG AAG	3168
	Ser Ala Gly Thr Val Phe Ser Phe Ala Pro Asp Met Gly Lys Ala Lys	
	1045 1050 1055	
5	AAT GCG GCC GCC GAA TTC CGA CGA CTG TTC GAC CGA AAG CCA CAA ATT	3216
	Asn Ala Ala Ala Gly Phe Arg Arg Leu Phe Asp Arg Lys Pro Gln Ile	
	1060 1065 1070	
10	GAT AAC TGG TCT GAA GAG GGC GAG AAG CTC GAA ACG GTG GAA GGT GAA	3264
	Asp Asn Trp Ser Glu Glu Gly Glu Lys Leu Glu Thr Val Glu Gly Glu	
	1075 1080 1085	
15	ATC GAA TTT AGG AAC GTG CAC TTC AGA TAC CCG ACC CGC CCA GAA CAG	3312
	Ile Glu Phe Arg Asn Val His Phe Arg Tyr Pro Thr Arg Pro Glu Gln	
	1090 1095 1100	
	CCT GTC CTG CGC GGC TTG GAC CTG ACC GTG AAG CCT GGA CAA TAT GTT	3360
	Pro Val Leu Arg Gly Leu Asp Leu Thr Val Lys Pro Gly Gln Tyr Val	
	1105 1110 1115 1120	
20	GCG CTT GTC GGA CCC AGC GGT TGT GGC AAG AGT ACC ACC ATT GCA TTG	3408
	Ala Leu Val Gly Pro Ser Gly Cys Gly Lys Ser Thr Thr Ile Ala Leu	
	1125 1130 1135	
25	CTT GAG CGC TTT TAC GAT GCG ATT GCC GGG TCC ATC CTT GTT GAT GGG	3456
	Leu Glu Arg Phe Tyr Asp Ala Ile Ala Gly Ser Ile Leu Val Asp Gly	
	1140 1145 1150	
30	AAG GAC ATA AGT AAA CTA AAT ATC AAC TCC TAC CGC AGC TTT CTG TCA	3504
	Lys Asp Ile Ser Lys Leu Asn Ile Asn Ser Tyr Arg Ser Phe Leu Ser	
	1155 1160 1165	
35	CTG GTC AGC CAG GAG CCG ACA CTG TAC CAG GGC ACC ATC AAG GAA AAC	3552
	Leu Val Ser Gln Glu Pro Thr Leu Tyr Gln Gly Thr Ile Lys Glu Asn	
	1170 1175 1180	
40	ATC TTA CTT GGT ATT GTC GAA GAT GAC GTA CCG GAA GAA TTC TTG ATT	3600
	Ile Leu Leu Gly Ile Val Glu Asp Asp Val Pro Glu Glu Phe Leu Ile	
	1185 1190 1195 1200	
	AAG GCT TGC AAG GAC GCT AAT ATC TAC GAC TTC ATC ATG TCG CTC CCG	3648
	Lys Ala Cys Lys Asp Ala Asn Ile Tyr Asp Phe Ile Met Ser Leu Pro	
	1205 1210 1215	
45	GAG GGC TTT AAT ACA GTT GTT GGC AGC AAG GGA GGC ATG TTG TCT GGC	3696
	Glu Gly Phe Asn Thr Val Val Gly Ser Lys Gly Gly Met Leu Ser Gly	
	1220 1225 1230	
50	GGC CAA AAG CAA CGT GTG GCC ATT GCC CGA GCC CTT CTT CGG GAT CCC	3744
	Gly Gln Lys Gln Arg Val Ala Ile Ala Arg Ala Leu Leu Arg Asp Pro	
	1235 1240 1245	
55	AAA ATC CTT CTT CTC GAT GAA GCG ACG TCA GCC CTC GAC TCC GAG TCA	3792
	Lys Ile Leu Leu Leu Asp Glu Ala Thr Ser Ala Leu Asp Ser Glu Ser	
	1250 1255 1260	
	GAA AAG GTC GTC CAG GCG GCT TTG GAT GCC GCT GCC CGA GGC CGA ACC	3840
	Glu Lys Val Val Gln Ala Ala Leu Asp Ala Ala Arg Gly Arg Thr	
	1265 1270 1275 1280	
60	ACA ATC GCC GTT GCA CAC CGA CTC AGC ACG ATT CAA AAG GCG GAC GTT	3888
	Thr Ile Ala Val Ala His Arg Leu Ser Thr Ile Gln Lys Ala Asp Val	
	1285 1290 1295	
65	ATC TAT GTT TTC GAC CAA GGC AAG ATC GTC GAA AGC GGA ACG CAC AGC	3936

Ile Tyr Val Phe Asp Gln Gly Lys Ile Val Glu Ser Gly Thr His Ser
 1300 1305 1310

5 GAA CTG GTC CAG AAA AAG GGC CGG TAC TAC GAG CTG GTC AAC TTG CAG 3984
 Glu Leu Val Gln Lys Lys Gly Arg Tyr Tyr Glu Leu Val Asn Leu Gln
 1315 1320 1325

AGC TTG GCC AAG GGC CAT 4002
 Ser Leu Gly Lys Gly His
 1330

10

(2) INFORMATION FOR SEQ ID NO:2:

15 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1334 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: protein

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:2:

25 Met Ser Pro Leu Glu Thr Asn Pro Leu Ser Pro Glu Thr Ala Met Arg
 1 5 10 15
 Glu Pro Ala Glu Thr Ser Thr Thr Glu Glu Gln Ala Ser Thr Pro His
 20 25 30

30 Ala Ala Asp Glu Lys Lys Ile Leu Ser Asp Leu Ser Ala Pro Ser Ser
 35 40 45
 Thr Thr Ala Thr Pro Ala Asp Lys Glu His Arg Pro Lys Ser Ser Ser
 50 55 60

35 Ser Asn Asn Ala Val Ser Val Asn Glu Val Asp Ala Leu Ile Ala His
 65 70 75 80

40 Leu Pro Glu Asp Glu Arg Gln Val Leu Lys Thr Gln Leu Glu Glu Ile
 85 90 95
 Lys Val Asn Ile Ser Phe Phe Gly Leu Trp Arg Tyr Ala Thr Lys Met
 100 105 110

45 Asp Ile Leu Ile Met Val Ile Ser Thr Ile Cys Ala Ile Ala Ala Ala
 115 120 125

Ser Thr Phe Gln Arg Ile Met Leu Tyr Gln Ile Ser Tyr Asp Glu Phe
 130 135 140

50 Tyr Asp Glu Leu Thr Lys Asn Val Leu Tyr Phe Val Tyr Leu Gly Ile
 145 150 155 160

55 Gly Glu Phe Val Thr Val Tyr Val Ser Thr Val Gly Phe Ile Tyr Thr
 165 170 175
 Gly Glu His Ala Thr Gln Lys Ile Arg Glu Tyr Tyr Leu Glu Ser Ile
 180 185 190

60 Leu Arg Gln Asn Ile Gly Tyr Phe Asp Lys Leu Gly Ala Gly Glu Val
 195 200 205

Thr Thr Arg Ile Thr Ala Asp Thr Asn Leu Ile Gln Asp Gly Ile Ser
 210 215 220

65

Glu Lys Val Gly Leu Thr Leu Thr Ala Leu Ala Thr Phe Val Thr Ala
 225 230 235
 5 Phe Ile Ile Ala Tyr Val Lys Tyr Trp Lys Leu Ala Leu Ile Cys Ser
 245 250 255
 Ser Thr Ile Val Ala Leu Val Leu Thr Met Gly Gly Gly Ser Gln Phe
 260 265 270
 10 Ile Ile Lys Tyr Ser Lys Lys Ser Leu Asp Ser Tyr Gly Ala Gly Gly
 275 280 285
 Thr Val Ala Glu Glu Val Ile Ser Ser Ile Arg Asn Ala Thr Ala Phe
 290 295 300
 15 Gly Thr Gln Asp Lys Leu Ala Lys Gln Tyr Glu Val His Leu Asp Glu
 305 310 315 320
 Ala Glu Lys Trp Gly Thr Lys Asn Gln Ile Val Met Gly Phe Met Ile
 325 330 335
 20 Gly Ala Met Phe Gly Leu Met Tyr Ser Asn Tyr Gly Leu Gly Phe Trp
 340 345 350
 25 Met Gly Ser Arg Phe Leu Val Asp Gly Ala Val Asp Val Gly Asp Ile
 355 360 365
 Leu Thr Val Leu Met Ala Ile Leu Ile Gly Ser Phe Ser Leu Gly Asn
 370 375 380
 30 Val Ser Pro Asn Ala Gln Ala Phe Thr Asn Ala Val Ala Ala Ala Ala
 385 390 395 400
 35 Lys Ile Phe Gly Thr Ile Asp Arg Gln Ser Pro Leu Asp Pro Tyr Ser
 405 410 415
 Asn Glu Gly Lys Thr Leu Asp His Phe Glu Gly His Ile Glu Leu Arg
 420 425 430
 40 Asn Val Lys His Ile Tyr Pro Ser Arg Pro Glu Val Thr Val Met Glu
 435 440 445
 Asp Val Ser Leu Ser Met Pro Ala Gly Lys Thr Thr Ala Leu Val Gly
 450 455 460
 45 Pro Ser Gly Ser Gly Lys Ser Thr Val Val Gly Leu Val Glu Arg Phe
 465 470 475 480
 Tyr Met Pro Val Arg Gly Thr Val Leu Leu Asp Gly His Asp Ile Lys
 485 490 495
 50 Asp Leu Asn Leu Arg Trp Leu Arg Gln Ile Ser Leu Val Ser Gln
 500 505 510
 55 Glu Pro Val Leu Phe Gly Thr Thr Ile Tyr Lys Asn Ile Arg His Gly
 515 520 525
 Leu Ile Gly Thr Lys Tyr Glu Asn Glu Ser Glu Asp Lys Val Arg Glu
 530 535 540
 60 Leu Ile Glu Asn Ala Ala Lys Met Ala Asn Ala His Asp Phe Ile Thr
 545 550 555 560
 Ala Leu Pro Glu Gly Tyr Glu Thr Asn Val Gly Gln Arg Gly Phe Leu
 565 570 575

Leu Ser Gly Gly Gln Lys Gln Arg Ile Ala Ile Ala Arg Ala Val Val
 580 585 590
 5 Ser Asp Pro Lys Ile Leu Leu Leu Asp Glu Ala Thr Ser Ala Leu Asp
 595 600 605
 Thr Lys Ser Glu Gly Val Val Gln Ala Ala Leu Glu Arg Ala Ala Glu
 610 615 620
 10 Gly Arg Thr Thr Ile Val Ile Ala His Arg Leu Ser Thr Ile Lys Thr
 625 630 635 640
 15 Ala His Asn Ile Val Val Leu Val Asn Gly Lys Ile Ala Glu Gln Gly
 645 650 655
 Thr His Asp Glu Leu Val Asp Arg Gly Gly Ala Tyr Arg Lys Leu Val
 660 665 670
 20 Glu Ala Gln Arg Ile Asn Glu Gln Lys Glu Ala Asp Ala Leu Glu Asp
 675 680 685
 Ala Asp Ala Glu Asp Leu Thr Asn Ala Asp Ile Ala Lys Ile Lys Thr
 690 695 700
 25 Ala Ser Ser Ala Ser Ser Asp Leu Asp Gly Lys Pro Thr Thr Ile Asp
 705 710 715 720
 30 Arg Thr Gly Thr His Lys Ser Val Ser Ser Ala Ile Leu Ser Lys Arg
 725 730 735
 Pro Pro Glu Thr Thr Pro Lys Tyr Ser Leu Trp Thr Leu Leu Lys Phe
 740 745 750
 35 Val Ala Ser Phe Asn Arg Pro Glu Ile Pro Tyr Met Leu Ile Gly Leu
 755 760 765
 Val Phe Ser Val Leu Ala Gly Gly Gly Gln Pro Thr Gln Ala Val Leu
 770 775 780
 40 Tyr Ala Lys Ala Ile Ser Thr Leu Ser Leu Pro Glu Ser Gln Tyr Ser
 785 790 795 800
 45 Lys Leu Arg His Asp Ala Asp Phe Trp Ser Leu Met Phe Phe Val Val
 805 810 815
 Gly Ile Ile Gln Phe Ile Thr Gln Ser Thr Asn Gly Ala Ala Phe Ala
 820 825 830
 50 Val Cys Ser Glu Arg Leu Ile Arg Arg Ala Arg Ser Thr Ala Phe Arg
 835 840 845
 Thr Ile Leu Arg Gln Asp Ile Ala Phe Phe Asp Lys Glu Glu Asn Ser
 850 855 860
 55 Thr Gly Ala Leu Thr Ser Phe Leu Ser Thr Glu Thr Lys His Leu Ser
 865 870 875 880
 60 Gly Val Ser Gly Val Thr Leu Gly Thr Ile Leu Met Thr Ser Thr Thr
 885 890 895
 Leu Gly Ala Ala Ile Ile Ile Ala Leu Ala Ile Gly Trp Lys Leu Ala
 900 905 910
 65 Leu Val Cys Ile Ser Val Val Pro Val Leu Leu Ala Cys Gly Phe Tyr

	915	920	925
	Arg Phe Tyr Met Leu Ala Gln Phe Gln Ser Arg Ser Lys Leu Ala Tyr		
5	930	935	940
	Glu Gly Ser Ala Asn Phe Ala Cys Glu Ala Thr Ser Ser Ile Arg Thr		
	945	950	955
10	Val Ala Ser Leu Thr Arg Glu Arg Asp Val Trp Glu Ile Tyr His Ala		
	965	970	975
	Gln Leu Asp Ala Gln Gly Arg Thr Ser Leu Ile Ser Val Leu Arg Ser		
	980	985	990
15	Ser Leu Leu Tyr Ala Ser Ser Gln Ala Leu Val Phe Phe Cys Val Ala		
	995	1000	1005
	Leu Gly Phe Trp Tyr Gly Gly Thr Leu Leu Gly His His Glu Tyr Asp		
	1010	1015	1020
20	Ile Phe Arg Phe Phe Val Cys Phe Ser Glu Ile Leu Phe Gly Ala Gln		
	1025	1030	1035
	Ser Ala Gly Thr Val Phe Ser Phe Ala Pro Asp Met Gly Lys Ala Lys		
	1045	1050	1055
25	Asn Ala Ala Ala Glu Phe Arg Arg Leu Phe Asp Arg Lys Pro Gln Ile		
	1060	1065	1070
30	Asp Asn Trp Ser Glu Glu Gly Glu Lys Leu Glu Thr Val Glu Gly Glu		
	1075	1080	1085
	Ile Glu Phe Arg Asn Val His Phe Arg Tyr Pro Thr Arg Pro Glu Gln		
	1090	1095	1100
35	Pro Val Leu Arg Gly Leu Asp Leu Thr Val Lys Pro Gly Gln Tyr Val		
	1105	1110	1115
	Ala Leu Val Gly Pro Ser Gly Cys Gly Lys Ser Thr Thr Ile Ala Leu		
	1125	1130	1135
40	Leu Glu Arg Phe Tyr Asp Ala Ile Ala Gly Ser Ile Leu Val Asp Gly		
	1140	1145	1150
45	Lys Asp Ile Ser Lys Leu Asn Ile Asn Ser Tyr Arg Ser Phe Leu Ser		
	1155	1160	1165
	Leu Val Ser Gln Glu Pro Thr Leu Tyr Gln Gly Thr Ile Lys Glu Asn		
	1170	1175	1180
50	Ile Leu Leu Gly Ile Val Glu Asp Asp Val Pro Glu Glu Phe Leu Ile		
	1185	1190	1195
	Lys Ala Cys Lys Asp Ala Asn Ile Tyr Asp Phe Ile Met Ser Leu Pro		
	1205	1210	1215
55	Glu Gly Phe Asn Thr Val Val Gly Ser Lys Gly Gly Met Leu Ser Gly		
	1220	1225	1230
60	Gly Gln Lys Gln Arg Val Ala Ile Ala Arg Ala Leu Leu Arg Asp Pro		
	1235	1240	1245
	Lys Ile Leu Leu Leu Asp Glu Ala Thr Ser Ala Leu Asp Ser Glu Ser		
	1250	1255	1260
65			

Glu Lys Val Val Gln Ala Ala Leu Asp Ala Ala Ala Arg Gly Arg Thr
 1265 1270 1275 1280
 5 Thr Ile Ala Val Ala His Arg Leu Ser Thr Ile Gln Lys Ala Asp Val
 1285 1290 1295
 Ile Tyr Val Phe Asp Gln Gly Lys Ile Val Glu Ser Gly Thr His Ser
 1300 1305 1310
 10 Glu Leu Val Gln Lys Lys Gly Arg Tyr Tyr Glu Leu Val Asn Leu Gln
 1315 1320 1325
 Ser Leu Gly Lys Gly His
 1330

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 4002 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (ii) MOLECULE TYPE: mRNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTI-SENSE: NO

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:3:

35	AUGUCCCCGC UAGAGACAAA UCCCCUUUCG CCAGAGACUG CUAUGCGCGA ACCUGCUGAG	60
	ACUUCAACGA CGGAGGAGCA AGCUUCUACA CCACACGUG CGGACGAGAA GAAAAUCCUC	120
40	AGCGACCUCU CGGCUCCAUC UAGUACUACA GCAACCCCGC CAGACAAGGA GCACCGUCCU	180
	AAAUUCGUCG CCAGCAAUAA UCGGUCUCUG GUCAACGAAG UCGAUGCGCU UAUUGCGCAC	240
	CUGCCAGAAG ACGAGAGGCA GGUCUUGAAG ACGCAGCUGG AGGAGAUCAA AGUAAACAU	300
45	UCCUUCUUCG GUCUCUGGCG GUUGCAACA AAGAUGGAUA UACUUAUCAU GGUAAUCAGU	360
	ACAAUUCUGG CCAUUGCUGC CGGUCGACU UUCCAGAGGA UAAUGUUAUA UCAAAUUCUG	420
	UACGACGAGU UCUAUGAUGA AUUGACCAAG AACGUACUGU ACUCUGUAUA CCUCGUAUC	480
50	GGCGAGUUUG UCACUGUCUA UGUUAGUACU GUUGGCUUCA UCUAUACCGG AGAACAAGCC	540
	ACGCAGAAGA UCCGCGAGUA UUACCUUGAG UCUAUCCUGC GCCAGAACA UGGCUAUUUU	600
55	GAUAAACUCG GUGCCGGGGA AGUGACCACC CGUAUAACAG CCGAUACAAA CCUUAUCCAG	660
	GAUGGCAUUU CGGAGAAGGU CGGUCUCACU UUGACUGCCC UGGCGACA AU CGUGACAGCA	720
60	UUCAUUAUCG CCUACGUCAA AUACUGGAAG UUGGCUUAA UUGUCAGCUC AACAAUUGUG	780
	GCCCCUGUUC UCACCAUGGG CGUGGUUUCU CAGUUUAUCA UCAAGUACAG CAAAAGUUC	840
	CUUGACAGCU ACGGUGCAGG CGGCACUGUU GCGGAAGAGG UCAUCAGCUC CAUCAGAAAU	900
65	GCCACAGCGU UUGGCACCCA AGACAAGCUU GCGAAGCAGU AUGAGGUCCA CUUAGACGAA	960

	GCUGAGAAAU	GGGGAACAAA	GAACCAGAUU	GUCAUGGGUU	UCAUGAUUGG	CGCCAUGUUU	1020
5	GGCCUUAUUG	ACUCGAACUA	CGGUCUUGGC	UUUCUGGAUG	GUUCUCGUUU	CCUGGUGAGU	1080
	GGUGCAAGUG	AUGUGGGUGA	UAUUCUACA	GUUCUCAUGG	CCACUUGAUU	CGGAUCGUUC	1140
	UCUUUGGGGA	ACGUUAGUCC	AAAUGCUCAA	GCAUUUACAA	ACGCUGUGGC	CGCGGCCGCA	1200
10	AAGAUUAUUU	GAACGAUCGA	UCGCCAGUCC	CCAUAAGAUC	CAUAUUCGAA	CGAAGGGAAG	1260
	ACGCUCGACC	AUUUUGAGGG	CCACAUGAG	UUACGCAAUG	UCAAGCAUUA	UUACCCAUUC	1320
15	AGACCCGAGG	UCACCGUCAU	GGAGGAUGUU	UCUCUGUCA	UGCCCGCUGG	AAAAACAACC	1380
	GCUUUAGUCG	GCCCCUCUGG	CUCUGGAAAA	AGUACGGUGG	UCGGCUUGGU	UGAGCGAUUC	1440
	UACAUGCCUG	UUCGCGGUAC	GGUUUUGCUG	GAUGGCCAUG	ACAUCAAGGA	CCUCAAUUCU	1500
20	CGCUGGCUUC	GCCAACAGAU	CUCUUUGGUU	AGCCAGGAGC	CUGUUUUUUU	UGGCACGACG	1560
	AUUUAUAAGA	AUAUUAGGCA	CGGUCUCAUC	GGCACAAGU	ACGAGAAUGA	AUCCGAGGAU	1620
25	AAGGUCGGG	AACUCAUCGA	GAACGCGGCA	AAAUUGGCGA	AUGCUCADGA	CUUUUAUACU	1680
	GCCUUGCCUG	AAGGUUAUGA	GACCAAUGUU	GGCAGCGUG	GCUUUCUCCU	UUCAGGUGGC	1740
	CAGAAACAGC	GCAUUGCAAU	CGCCCCGUGC	GUUGUAGUG	ACCCAAAAAU	CCUGCUCUG	1800
30	GAUGAAGCUA	CUUCGGCCUU	GGACACAAAA	UCCGAAGGCG	UGGUUCAAAG	AGCUUUGGAG	1860
	AGGGCAGCUG	AAGGCCGAAC	UACUAUUGUG	AUCGCUCAUC	GCCUUUCCAC	GAUCAAACAG	1920
35	GCGCACACAA	UUGUGGUUCU	GGUCAAUUGC	AAAAUUGCUG	AACAAGGAAC	UCAAGAUCAA	1980
	UUGGUUGACC	GCGGAGGCGC	UUUUCGCAAA	CUUGUGGAGG	CUCAACGUUU	CAAUGAACAG	2040
	AAGGAAGCUG	ACGCCUUGGA	GGACGCCGAC	GCUGAGGAUC	UACGAAUGC	AGAUAUUGCC	2100
40	AAAAUCAAAA	CUGCGUCAAG	CGCAUCAUCC	GAUUCGACG	GAAAACCCAC	AACCAUUGAC	2160
	CGCACGGGCA	CCACAAGUUC	UGUUUCCAGC	GCGAUUCUUU	CUAAAAGACC	CCCCGAAACA	2220
45	ACUCCGAAAU	ACUUAUUAUG	GACGCGUCUC	AAAUUUGUUG	CUUCCUCAA	CCGCCUGGAA	2280
	AUCCCGUACA	UGCUCUACGG	UCUUGUCUUC	UCAGUGUUAU	CUGGUGGUGG	CCAACCCACG	2340
	CAAGCAGUGC	UAUAUGCUAA	AGCCAUACAG	ACACUCUCCG	UCCCAGAAUC	ACAAUAUAGC	2400
50	AAGCUUCGAC	AUGAUGCGGA	UUUCUGGUCA	UUGAUUUUCU	UCGUGGUUGG	UAUCAUUCAG	2460
	UUUAUCACGC	AGUCAACCAA	UGGUGUGUCA	UUUGCCGUUU	GCUCGAGAG	ACUUAUUCUG	2520
55	CGCGCGAGAA	GCACUGCCUU	UCGGACGAUA	CUCCGUCAG	ACAUUGCUUU	CUUUAGCAAG	2580
	GAAGAGAUAU	GCACCGCGCG	UCUGACCUUCU	UUUCUGUCCA	CCGAGACGAA	GCAUCUUCUC	2640
	GGUUAUAGCG	GUGUGACUCU	AGGCACGAUC	UUGAUGACCU	CCACGACCCU	AGGAGCGGCU	2700
60	AUCAUAUUG	CCUUGGCGAU	UGGGUGGAAA	UUGGCCUUAG	UUUGUAUUCU	GGUUGUGCCG	2760
	GUUCUCCUGG	CAUGCGGUUU	CUACCGAUUC	UAUAUGCUAG	CCCAGUUUCA	AUCACGCUCC	2820
65	AAGCUUGCUU	AUGAGGGAUC	UGCAAAACUUU	GCUUGCGAGG	CUACAUCGUC	UAUCCGCACA	2880

	GUUGCGUCAU	UAACCCGGGA	AAGGGAUGUC	UGGAGAUUU	ACCAUGCCCA	GCUGACGCA	2940
	CAAGGCAGGA	CCAGUCUAAU	CUCUGUCUUG	AGGUCAUCC	UGUUUAUAGC	GUCGUCGAG	3000
5	GCACUUGUUU	UCUUCUGCGU	UGCUCUCGGG	UUUUGGUACG	GAGGGACACU	UCUUGGUCAC	3060
	CACGAGUAUG	ACAUUUUCCG	CUUCUUUGUU	UGUUUCUCCG	AGAUUCUCUU	UGGUCUCUAA	3120
	UCCGCGGGCA	CCGUCUUUUC	CUUUGCACCA	GACAUGGGCA	AGGCGAAGAA	UGC GCCGCC	3180
10	GAAUUCGAC	GACUGUUCGA	CCGAAAGCCA	CAAAUUGAUA	ACUGGUCUGA	AGAGGGCGAG	3240
	AAGCUCGAAA	CGGUGGAAGG	UGAAAUCGAA	UUUAGGAACG	UGCACUUCAG	AUACCCGACC	3300
15	CGCCAGAAAC	AGCCUGUCCU	GCGCGGCUUG	GACCUGACCG	UGAAGCCUGG	ACAAUAUGUU	3360
	GCGCUUGUCG	GACCCAGCGG	UUGUGGCAAG	AGUACCACCA	UUGCAUUGCU	UGAGCGCUUU	3420
	UACGAUGCGA	UUGCCGGGUC	CAUCCUUGUU	GAUGGGAAGG	ACAAUAGUAA	ACUAAAUAUC	3480
20	AACUCCUACC	GCAGCUUUCU	GUCACUGGUC	AGCCAGGAGC	CGACACUGUA	CCAGGGCACC	3540
	AUCAAGAAA	ACAUCUUAUC	UGGUAUUGUC	GAAGAUGACG	UACCGGAAGA	AUUCUUGAUU	3600
25	AAGGCUUGCA	AGGACGCUAA	UAUCUACGAC	UUCAUCAUGU	CGCUCCCGGA	GGGCUUUAUU	3660
	ACAGUUGUUG	GCAGCAAGGG	AGGCAUGUUG	UCUGCGGGCC	AAAAGCAACG	UGUGGCCAUU	3720
	GCCGAGGCC	UUCUUCGGGA	UCCCAAAUUC	CUUCUUCUCG	AUGAAGCGAC	GUCAGCCUUC	3780
30	GACUCCGAGU	CAGAAAAGGU	CGUCCAGGCG	GCUUUGGAUG	CCGUCUGCCG	AGGCCGAACC	3840
	ACAAUCGCGG	UUGCACACCG	ACUCAGCACG	AUUCAAAAGG	CGSACGUUAU	CUAUGUUUUC	3900
35	GACCAAGGCA	AGAUCGUCGA	AAGCGGAACG	CACAGCGAAC	UGGUCCAGAA	AAAGGGCCGG	3960
	UACUACGAGC	UGGUCACUUC	GCAGAGCUUG	GGCAAGGGCC	AU		4002